

CRLB Based Optimal Noise Enhanced Parameter Estimation Using Quantized Observations

Gökçe Osman Balkan, *Student Member, IEEE*, and Sinan Gezici, *Member, IEEE*

Abstract—In this letter, optimal additive noise is characterized for parameter estimation based on quantized observations. First, optimal probability distribution of noise that should be added to observations is formulated in terms of a Cramer–Rao lower bound (CRLB) minimization problem. Then, it is proven that optimal additive “noise” can be represented by a constant signal level, which means that randomization of additive signal levels is not needed for CRLB minimization. In addition, the results are extended to the cases in which there exists prior information about the unknown parameter and the aim is to minimize the Bayesian CRLB (BCRLB). Finally, a numerical example is presented to explain the theoretical results.

Index Terms—Cramer–Rao lower bound, estimation, noise enhanced estimation, quantization.

I. INTRODUCTION

ALTHOUGH noise commonly degrades the performance of a system, some nonlinear systems can benefit from addition of noise to their inputs or from increased noise levels [1], [2]. Advantages of additive noise are investigated also for parameter estimation problems. In the frequency estimation problem studied in [3], it is observed that, under certain conditions, the mean-squared error (MSE) of the optimal Bayesian estimator can reduce when the noise level is raised. Similarly, [4] considers Bayesian estimation and provides examples of when increased noise levels result in improved MSE performance. In [3] and [4], 1-bit quantizers are employed and noise benefits are observed due to the nonlinear structure of the quantizers. In addition, [5] studies parameter estimation based on 1-bit dithered quantization and proposes an estimator that does not require any information about the dither signal and the noise distribution. In another noise enhanced estimation study [6], the first and the second moments of an estimator and a Bayesian cost function are used as performance criteria and the general form of the optimal noise probability density function (p.d.f.) is derived.

For some noise enhanced parameter estimation problems, asymptotical behaviors of the estimators make the Cramer–Rao lower bound (CRLB), equivalently the Fisher information, an appealing metric for the quantification of performance improvements via additive noise. For example, maximization of the Fisher information for parameter estimation based on

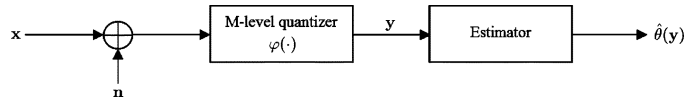


Fig. 1. Block diagram of the system, where \mathbf{n} denotes the additive noise that is independent of the original observation \mathbf{x} .

quantized observations is studied in [7] by optimizing quantization intervals. In addition, the dependence of the MSE of a mean estimator on the probability distribution of observation noise is investigated in [8] and theoretical lower bounds are provided. In [9], parameter estimation based on observations from a multi-bit quantizer is considered and additive controlled perturbation of the quantizer thresholds is investigated. In particular, [9] shows that random dithering can significantly reduce the CRLB for the mean estimation problem with 1-bit precision sampling. Moreover, it is shown in [10] that the variance of an estimator that uses 1-bit quantizer outputs can be made quite close to the variance of a clairvoyant estimator that uses unquantized observations by an appropriate choice of the quantizer threshold. Furthermore, addition of noise to quantized measurements can provide enhancement of the Fisher information for the estimation of the suprathreshold input signals [11]. Finally, maximization of the Fisher information by both an appropriate choice of the quantizer threshold and additive noise is studied in [12].

Although the effects of additive noise on CRLBs have been investigated in [9], [11] and [12], the optimal p.d.f. of additive noise that minimizes the CRLB for parameter estimation based on quantized observations has not been obtained before. In this letter, a parameter estimation problem based on quantized observations is studied, where the aim is to find the optimal p.d.f. of noise that should be added to the observations before the quantizer in order to minimize the CRLB for estimating the unknown parameter (cf. Fig. 1). Unlike the previous studies, an explicit CRLB minimization problem is formulated in terms of the additive noise p.d.f., the quantization function, and the p.d.f. of the original observation. It is first shown, for a given value of the parameter, that optimal additive “noise” can be represented by a specific constant value. In addition, the results are extended to the cases in which there exists prior information about the unknown parameter, and it is shown that the Bayesian Cramer–Rao lower bound (BCRLB) is minimized when the additive “noise” is represented by a constant value.

II. PROBLEM FORMULATION

Consider a system in which a quantized version of observation \mathbf{x} is used to estimate an underlying parameter θ . Let $p_{\mathbf{X}}(\mathbf{x}; \theta)$ represent the p.d.f. of the observation, and $\varphi(\cdot)$ denote the quantizer. Instead of using observation \mathbf{x} , a noise modified version of the observation, $\mathbf{x} + \mathbf{n}$, can be used as in Fig. 1 in order to improve the estimation accuracy of the system, where

Manuscript received December 22, 2009; revised February 04, 2010. First published February 22, 2010; current version published March 31, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Aleksandar Dogandzic.

The authors are with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey (e-mails: balkango@ee.bilkent.edu.tr; gezici@ee.bilkent.edu.tr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2010.2043787

the additive noise \mathbf{n} is independent of the observation \mathbf{x} . The aim is to obtain the p.d.f. of \mathbf{n} , denoted by $p_{\mathbf{N}}(\cdot)$, that maximizes the estimation accuracy of the system in Fig. 1. It is noted that this noise enhanced parameter estimation problem can also be regarded as a dynamic bias control problem as in [9], when \mathbf{n} represents the control input for the quantizer bias.

Suppose that quantizer $\varphi(\cdot)$ is an M -level quantizer that generates the quantized observation vector \mathbf{y} based on the noise modified input observation as follows:

$$\mathbf{y} = \varphi(\mathbf{x} + \mathbf{n}) \quad (1)$$

where $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_L]$, $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_L]$, $\mathbf{n} = [n_1 \ n_2 \ \dots \ n_L]$, and the quantizer levels are determined by thresholds $\tau_1, \tau_2, \dots, \tau_{M-1}$. Specifically, the relation between the input and the output of the quantizer is described, for $l = 1, 2, \dots, L$, as

$$y_l = i, \quad \text{if } \tau_i < x_l + n_l \leq \tau_{i+1} \quad (2)$$

for $i = 0, 1, \dots, M-1$, where $\tau_0 \triangleq -\infty$ and $\tau_M \triangleq \infty$ are used for the simplicity of the expressions.

Let $p_{\mathbf{Y}}(\cdot; \theta)$ represent the probability mass function (p.m.f.) of the quantizer output for a given value of θ . From (2), it can be obtained as

$$p_{\mathbf{Y}}(\mathbf{i}; \theta) = \int_{\mathbb{R}^L} P(\tau_{i_1} - n_1 < X_1 \leq \tau_{i_1+1} - n_1, \dots, \tau_{i_L} - n_L < X_L \leq \tau_{i_L+1} - n_L) p_{\mathbf{N}}(\mathbf{n}) d\mathbf{n} \quad (3)$$

for $\mathbf{i} \in \mathcal{I} \triangleq \{0, 1, \dots, M-1\}^L$, where i_l represents the l th component of \mathbf{i} .

The additive noise component \mathbf{n} in Fig. 1 is optimized according to the CRLB in this study; that is, the optimal noise p.d.f. that minimizes the CRLB is sought for. The CRLB on the MSE of unbiased estimators $\hat{\theta}$ of θ is stated as

$$\text{MSE}_{\theta}\{\hat{\theta}\} \geq J_{\theta}^{-1} = \left(\mathbb{E} \left\{ \left(\frac{\partial \log p_{\mathbf{Y}}(\mathbf{y}; \theta)}{\partial \theta} \right)^2 \right\} \right)^{-1} \quad (4)$$

where $\text{MSE}_{\theta}\{\hat{\theta}\} = \mathbb{E}\{(\hat{\theta}(\mathbf{y}) - \theta)^2\}$, J_{θ} is defined as the Fisher information [13], and $p_{\mathbf{Y}}(\cdot; \theta)$ is as in (3). Since the CRLB imposes a lower limit on the MSE of an unbiased estimator and since some estimators, such as the maximum likelihood estimator, can (asymptotically) achieve the CRLB under certain conditions [13], the aim in this study is to obtain the optimal p.d.f. of the additive noise that minimizes the CRLB specified by (4).

As the CRLB is the inverse of the Fisher information, the optimal additive noise p.d.f. can be formulated, from (4), as the solution of the following optimization problem:

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \max_{p_{\mathbf{N}}(\cdot)} \mathbb{E} \left\{ \left(\frac{\partial \log p_{\mathbf{Y}}(\mathbf{y}; \theta)}{\partial \theta} \right)^2 \right\}. \quad (5)$$

Since \mathbf{Y} is equal to \mathbf{i} with probability $p_{\mathbf{Y}}(\mathbf{i}; \theta)$ as defined in (3), the problem in (5) can be expressed as

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \max_{p_{\mathbf{N}}(\cdot)} \sum_{\mathbf{i} \in \mathcal{I}} \frac{1}{p_{\mathbf{Y}}(\mathbf{i}; \theta)} \left(\frac{\partial p_{\mathbf{Y}}(\mathbf{i}; \theta)}{\partial \theta} \right)^2. \quad (6)$$

As a special case of the generic problem formulation in (6), when both \mathbf{X} and \mathbf{N} consist of independent components, it can be shown that the components of the optimal additive noise can be calculated separately; i.e.,

$$p_{N_l}^{\text{opt}}(n) = \arg \max_{p_{N_l}(\cdot)} \sum_{i=0}^{M-1} \frac{1}{p_{Y_l}(i; \theta)} \left(\frac{\partial p_{Y_l}(i; \theta)}{\partial \theta} \right)^2 \quad (7)$$

for $l = 1, \dots, L$, where $p_{N_l}(\cdot)$ represents the marginal p.d.f. of the l th component of the additive noise, and $p_{Y_l}(i; \theta)$ denotes the probability that Y_l is equal to i for $i = 0, 1, \dots, M-1$. In addition, if Y_1, \dots, Y_L are independent and identically distributed (i.i.d.); that is, if $p_{Y_l}(i; \theta) = p_Y(i; \theta)$ for $l = 1, \dots, L$, the optimization problems in (7) become identical.

III. STATISTICAL CHARACTERIZATION OF OPTIMAL ADDITIVE NOISE

In order to investigate the statistical properties of the optimal additive noise in (6), we first introduce

$$H_{\mathbf{i}}^{\theta}(\mathbf{n}) \triangleq P(\tau_{i_1} - n_1 < X_1 \leq \tau_{i_1+1} - n_1, \dots, \tau_{i_L} - n_L < X_L \leq \tau_{i_L+1} - n_L), \quad (8)$$

$$G_{\mathbf{i}}^{\theta}(\mathbf{n}) \triangleq \frac{\partial H_{\mathbf{i}}^{\theta}(\mathbf{n})}{\partial \theta}. \quad (9)$$

It is noted from (3) that $0 \leq H_{\mathbf{i}}^{\theta}(\mathbf{n}) \leq 1$, $\forall \mathbf{n}$, and that $\sum_{\mathbf{i} \in \mathcal{I}} H_{\mathbf{i}}^{\theta}(\mathbf{n}) = 1$. Based on the definitions in (8) and (9), the p.m.f. in (3) and its derivative with respect to θ can be expressed as $p_{\mathbf{Y}}(\mathbf{i}; \theta) = \mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}$ and $\partial p_{\mathbf{Y}}(\mathbf{i}; \theta)/\partial \theta = \mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\}$. Then, the optimization problem in (6) becomes

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \max_{p_{\mathbf{N}}(\cdot)} \sum_{\mathbf{i} \in \mathcal{I}} \frac{(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2}{\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}}. \quad (10)$$

In order to obtain the solution of (10), the following lemma is presented first.

Lemma 1: For the real-valued functions defined in (8) and (9),

$$\sum_{\mathbf{i} \in \mathcal{I}} \frac{(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2}{\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}} \leq \max_{\mathbf{n}} \left\{ \sum_{\mathbf{i} \in \mathcal{I}} \frac{(G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2}{H_{\mathbf{i}}^{\theta}(\mathbf{n})} \right\} \quad (11)$$

is satisfied for all θ and all possible p.d.f.s $p_{\mathbf{N}}(\cdot)$ of \mathbf{N} .

Proof: ¹Consider a function of two variables defined as $f(\mathbf{Z}) = Z_1^2/Z_2$, where $\mathbf{Z} = [Z_1 \ Z_2]$. After some manipulation, the Hessian of $f(\mathbf{Z})$ can be shown to be positive semidefinite; hence, $f(\mathbf{Z})$ is convex, for $Z_2 \geq 0$. Therefore, Jensen's inequality implies that $(\mathbb{E}\{Z_1\})^2/\mathbb{E}\{Z_2\} \leq \mathbb{E}\{Z_1^2/Z_2\}$, for $Z_2 \geq 0$, which, upon the definition of $Z_1 \triangleq G_{\mathbf{i}}^{\theta}(\mathbf{N})$ and $Z_2 \triangleq H_{\mathbf{i}}^{\theta}(\mathbf{N})$, becomes $(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2/\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\} \leq \mathbb{E}\{(G_{\mathbf{i}}^{\theta}(\mathbf{N}))^2/H_{\mathbf{i}}^{\theta}(\mathbf{N})\}$ for all $p_{\mathbf{N}}(\cdot)$, θ and \mathbf{i} , since $H_{\mathbf{i}}^{\theta}(\mathbf{n}) \geq 0$, $\forall \mathbf{n}, \mathbf{i}, \theta$, by definition (cf. (8)). As this inequality is valid for all \mathbf{i} 's, we obtain

$$\sum_{\mathbf{i} \in \mathcal{I}} \frac{(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2}{\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}} \leq \mathbb{E} \left\{ \sum_{\mathbf{i} \in \mathcal{I}} \frac{(G_{\mathbf{i}}^{\theta}(\mathbf{N}))^2}{H_{\mathbf{i}}^{\theta}(\mathbf{N})} \right\} \quad (12)$$

for all $p_{\mathbf{N}}(\cdot)$ and θ . Finally, as the expression on the right-hand-side of (12) is never larger than $\max_{\mathbf{n}} \{\sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2/H_{\mathbf{i}}^{\theta}(\mathbf{n})\}$, the result in the lemma is obtained. \square

Lemma 1 states that for each possible noise p.d.f. $p_{\mathbf{N}}(\mathbf{n})$, the Fisher information $\sum_{\mathbf{i} \in \mathcal{I}} (\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2/\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}$ can never be larger than the maximum of $\sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2/H_{\mathbf{i}}^{\theta}(\mathbf{n})$ over all possible noise values, \mathbf{n} . In other words, Lemma 1 states

¹The authors thank the reviewer who suggested the approach in the proof.

that randomization among different noise values cannot improve (increase) the objective function in (10). This result leads to the following proposition.

Proposition 1: The optimal noise p.d.f. in (10) can be expressed as $p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$, where $\mathbf{n}_o = \arg \max_{\mathbf{n}} \sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2 / H_{\mathbf{i}}^{\theta}(\mathbf{n})$.

Proof: Since the result in Lemma 1 holds for any $p_{\mathbf{N}}(\cdot)$, the following inequality can be obtained:

$$\max_{p_{\mathbf{N}}(\cdot)} \left\{ \sum_{\mathbf{i} \in \mathcal{I}} \frac{(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2}{\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}} \right\} \leq \max_{\mathbf{n}} \left\{ \sum_{\mathbf{i} \in \mathcal{I}} \frac{(G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2}{H_{\mathbf{i}}^{\theta}(\mathbf{n})} \right\}. \quad (13)$$

Therefore, the maximum value of the objective function in (10) can never be larger than the expression on the right-hand-side of (13). However, this upper bound is achievable for $p_{\mathbf{N}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$, where \mathbf{n}_o is defined as in the proposition. Hence, the optimal additive noise can be expressed as in the proposition. \square

Proposition 1 states that for any additive noise that has a p.d.f. with multiple mass points, there always exists a corresponding constant “noise” level that provides an equal or smaller CRLB. In addition, it is noted from Lemma 1 and Proposition 1 that a constant additive “noise” component is optimal irrespective of the number of quantization levels (M) and the dimension of the observation vector (L). In addition, no assumption is imposed on the p.d.f. of the original observation, \mathbf{x} .

IV. OPTIMAL ADDITIVE NOISE IN THE PRESENCE OF PRIOR INFORMATION

In Section III, the optimal additive noise is calculated for a given value of θ . Although the value of θ is unknown in practice, the theoretical analysis in the previous section is useful in two aspects. First, it provides theoretical performance limits for unbiased estimators that perform parameter estimation based on quantized observations. Second, the theoretical results in the previous section form a basis for more practical results, and the ideas can be extended to the cases of unknown parameters. In the following, it is assumed that the exact value of θ is unknown, but its p.d.f., denoted by $w(\theta)$, is known *a priori*. Then, it is shown that the results in Lemma 1 and Proposition 1 can be extended to characterize the optimal additive noise.

In the presence of prior p.d.f. $w(\theta)$ for the unknown parameter θ , the Bayesian CRLB (BCRLB), also known as the posterior CRLB [14], imposes a lower bound on the MSE of any estimator $\hat{\theta}$ as [13]

$$\text{MSE}\{\hat{\theta}\} = \mathbb{E} \left\{ \left(\hat{\theta}(\mathbf{y}) - \theta \right)^2 \right\} \geq (J_D + J_P)^{-1} \quad (14)$$

where J_D and J_P represent the information obtained from the data (observations) and from the prior knowledge, respectively, and are given by

$$J_D = \mathbb{E} \left\{ \left(\frac{\partial \log p_{\mathbf{Y}}(\mathbf{y}; \theta)}{\partial \theta} \right)^2 \right\}, \quad J_P = \mathbb{E} \left\{ \left(\frac{\partial \log w(\theta)}{\partial \theta} \right)^2 \right\}. \quad (15)$$

It is important to note that the expectation in (14) is over both \mathbf{y} and θ , whereas that in (4) is over \mathbf{y} only.

Since J_P depends only on the prior p.d.f., it is independent of the additive noise component. Therefore, the optimal additive noise p.d.f. is defined to be the one that maximizes J_D . Then, similar to (5) and (6), the optimal additive noise p.d.f. can be formulated as

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \max_{p_{\mathbf{N}}(\cdot)} \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} \frac{1}{p_{\mathbf{Y}}(\mathbf{i}; \theta)} \left(\frac{\partial p_{\mathbf{Y}}(\mathbf{i}; \theta)}{\partial \theta} \right)^2 d\theta. \quad (16)$$

In other words, the aim now becomes maximizing the average of Fisher information J_{θ} [cf. (4)–(6)] for different parameter values. Since $p_{\mathbf{Y}}(\mathbf{i}; \theta) = \mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}$ and $\partial p_{\mathbf{Y}}(\mathbf{i}; \theta) / \partial \theta = \mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\}$ as defined in Section III, (16) can also be expressed as

$$p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \arg \max_{p_{\mathbf{N}}(\cdot)} \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} \frac{(\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2}{\mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\}} d\theta. \quad (17)$$

Then, the following proposition presents the p.d.f. of the optimal additive noise.

Proposition 2: The optimal noise p.d.f. in (17) can be expressed as $p_{\mathbf{N}}^{\text{opt}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$, where

$$\mathbf{n}_o = \arg \max_{\mathbf{n}} \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} \frac{(G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2}{H_{\mathbf{i}}^{\theta}(\mathbf{n})} d\theta. \quad (18)$$

Proof: Consider the inequality in (12), which is valid for all θ and $p_{\mathbf{N}}(\cdot)$. Since it holds for all θ values, the following inequality can be obtained: $\int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} (\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2 / \mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\} d\theta \leq \mathbb{E}\{\int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{N}))^2 / H_{\mathbf{i}}^{\theta}(\mathbf{N}) d\theta\}$ for all $p_{\mathbf{N}}(\cdot)$. Therefore, the maximum value of the objective function in (17) can be bounded from above by

$$\max_{p_{\mathbf{N}}(\cdot)} \mathbb{E} \left\{ \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} \frac{(G_{\mathbf{i}}^{\theta}(\mathbf{N}))^2}{H_{\mathbf{i}}^{\theta}(\mathbf{N})} d\theta \right\}. \quad (19)$$

Since the upper bound in (19) is always smaller than or equal to $\max_{\mathbf{n}} \{\int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{n}))^2 / H_{\mathbf{i}}^{\theta}(\mathbf{n}) d\theta\}$, the inequality $\max_{p_{\mathbf{N}}(\cdot)} \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} (\mathbb{E}\{G_{\mathbf{i}}^{\theta}(\mathbf{N})\})^2 / \mathbb{E}\{H_{\mathbf{i}}^{\theta}(\mathbf{N})\} d\theta \leq \int w(\theta) \sum_{\mathbf{i} \in \mathcal{I}} (G_{\mathbf{i}}^{\theta}(\mathbf{n}_o))^2 / H_{\mathbf{i}}^{\theta}(\mathbf{n}_o) d\theta$ can be obtained, where \mathbf{n}_o is as defined in (18). Since the upper bound in this inequality can be achieved for $p_{\mathbf{N}}(\mathbf{n}) = \delta(\mathbf{n} - \mathbf{n}_o)$, the result in the proposition is obtained. \square

Proposition 2 states that among all possible p.d.f.s for the additive noise components, a p.d.f. with a single mass point (that is, a constant “noise” component) minimizes the BCRLB. Therefore, adding the optimal noise to the observation is equivalent to shifting the threshold levels of the quantizer, which is a simple operation since no randomization among different noise values is needed.

V. NUMERICAL RESULTS AND CONCLUSIONS

In order to provide an example of the results in the previous sections, consider a scalar observation x in Fig. 1 with a Gaussian mixture p.d.f. given by $p_X(x; \theta) = 0.5\gamma(x; -\theta, \sigma^2) + 0.5\gamma(x; \theta, \sigma^2)$, where $\gamma(x; \theta, \sigma^2) \triangleq \exp\{-(x - \theta)^2 / (2\sigma^2)\} / (\sqrt{2\pi}\sigma)$. Then, $H_i^{\theta}(n)$ in (8) can be expressed as $H_i^{\theta}(n) = F_X(\tau_{i+1} - n; \theta) - F_X(\tau_i - n; \theta)$, for $i = 0, 1, \dots, M - 1$, where the cumulative distribution function (c.d.f.) of X for a given value of θ is calculated as $F_X(n; \theta) = 0.5Q(-(n - \theta)/\sigma) + 0.5Q(-(n + \theta)/\sigma)$, with $Q(a) = (2\pi)^{-0.5} \int_a^{\infty} e^{-0.5t^2} dt$ denoting the Q -function. Also, $G_i^{\theta}(n)$ in (9) can be calculated as the derivative of $H_i^{\theta}(n)$ with respect to θ . In addition, the quantizer in (2) is modeled

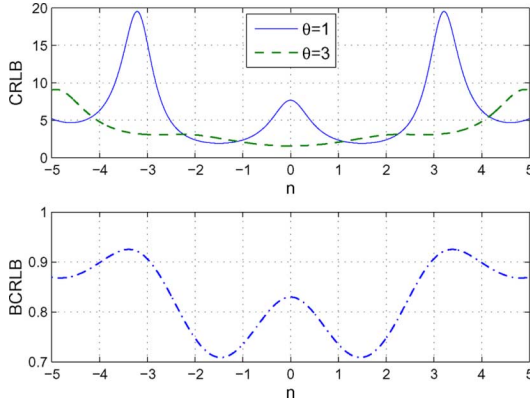


Fig. 2. (Top) CRLB versus additive “noise” n for $\theta = 1$ and $\theta = 3$. (Bottom) BCRLB versus n when θ is Gaussian distributed with unit mean and variance.

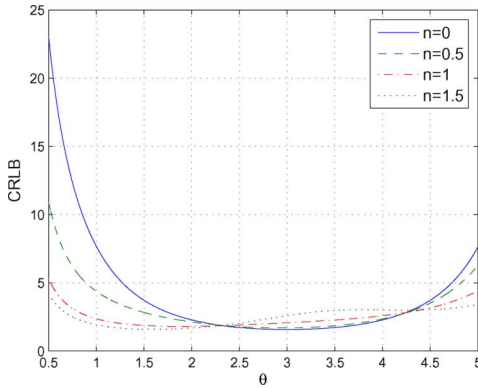


Fig. 3. CRLB versus θ for various values of additive “noise” n .

as a 4-level quantizer (i.e., $M = 4$) specified by thresholds $\tau_1 = -3$, $\tau_2 = 0$ and $\tau_3 = 3$.

First, optimal additive noise is investigated for given values of θ . Since Proposition 1 states that optimal additive noise n in Fig. 1 can be represented by a constant, the top plot in Fig. 2 investigates the CRLB versus constant “noise” levels for $\theta = 1$ and $\theta = 3$, where $\sigma = 1$ is used. Specifically, the inverse of the objective function in (10) is plotted against the additive “noise” level, n . It is observed for $\theta = 3$ that the optimal additive “noise” value is equal to zero, which means that the additive “noise” cannot reduce the CRLB of the system in that case. However, for $\theta = 1$, the minimum CRLB is achieved for $n = \pm 1.496$, which shows that additive “noise” n can result in a smaller CRLB. In addition, Fig. 3 plots the CRLB versus θ for various values of the additive “noise,” n . It is observed that the minimum CRLBs can be achieved by different n values over different ranges of parameter θ .

Next, for the problem setting described above, it is assumed that the prior p.d.f. of θ is specified as $w(\theta) = \exp\{-(\theta - \mu_\theta)^2 / (2\sigma_\theta^2)\} / (\sqrt{2\pi}\sigma_\theta)$, where $\mu_\theta = \sigma_\theta = 1$. From (15), it can be shown that $J_P = \sigma_\theta^{-2} = 1$. In Fig. 2, the BCRLB is plotted versus n , where the BCRLB is calculated as $(J_P + J_D)^{-1}$, with $J_P = 1$ and J_D denoting the value of the objective function in (18) for various values of n . It is observed from the figure that the minimum BCRLB is achieved at $n = \pm 1.46$. In addition, since there exists prior information in this scenario, the theoretical limits are lower than those in the previous scenario in which no prior information on θ exists.

TABLE I
UNIFORM DITHERING VERSUS OPTIMAL ADDITIVE “NOISE”

	$\epsilon = 1$	$\epsilon = 0.5$	$\epsilon = 0.25$	$\epsilon = 0$	Optimal
CRLB ($\theta = 1$)	6.566	7.302	7.575	7.675	1.924
CRLB ($\theta = 3$)	2.146	1.705	1.604	1.571	1.571
BCRLB	0.842	0.833	0.831	0.830	0.709

Finally, for the scenarios in Fig. 2, the lower bounds achieved by optimal additive noise are compared to those obtained via uniform dithering, which employs uniform additive noise between $-\epsilon$ and ϵ . The results in Table I reveal that uniform dithering can result in larger BCRLB values. For a given θ , the CRLB can decrease or increase with ϵ , as observed for $\theta = 1$ and $\theta = 3$, respectively. In all cases, uniform dithering can never achieve smaller lower bounds than the optimal additive noise in accordance with the theoretical results.

Since Propositions 1 and 2 state that optimal additive “noise” can be represented by a constant signal level, it is concluded that the CRLB (BCRLB) is minimized by shifting the original observation, which corresponds to shifting the thresholds of the quantizer by a constant value (cf. (2)). That is, among all possible p.d.f.s for the additive noise in Fig. 1, the ones with a single mass point, i.e., constant “noise” levels, can be used to achieve the minimum CRLB (BCRLB).

REFERENCES

- [1] R. Benzi, A. Sutera, and A. Vulpiani, “The mechanism of stochastic resonance,” *J. Phys. A: Math. Gen.*, vol. 14, pp. 453–457, 1981.
- [2] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, “Stochastic resonance,” *Rev. Mod. Phys.*, vol. 70, pp. 223–287, Jan. 1998.
- [3] F. Chapeau-Blondeau and D. Rousseau, “Noise-enhanced performance for an optimal Bayesian estimator,” *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1327–1334, May 2004.
- [4] F. Chapeau-Blondeau, “Noise-aided nonlinear Bayesian estimation,” *Phys. Rev. E*, vol. 66, no. 3, pp. 1–3, Sep. 2002.
- [5] O. Dabeer and A. Karnik, “Signal parameter estimation using 1-bit dithered quantization,” *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5389–5405, Dec. 2006.
- [6] H. Chen, P. K. Varshney, and J. H. Michels, “Noise enhanced parameter estimation,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.
- [7] S. Marano, V. Matta, and P. Willett, “Quantizer precision for distributed estimation in a large sensor network,” *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 4073–4078, Oct. 2006.
- [8] M. D. McDonnell, “Signal estimation via averaging of coarsely quantized signals,” in *Proc. Information, Decision and Control (IDC)*, Adelaide, Australia, Feb. 2007, pp. 100–105.
- [9] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim, “Sequential signal encoding from noisy measurements using quantizers with dynamic bias control,” *IEEE Trans. Inf. Theory*, vol. 47, pp. 978–1002, 2001.
- [10] A. Ribeiro and G. B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks—Part II: Unknown probability density function,” *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, Jul. 2006.
- [11] D. Rousseau, F. Duan, and F. Chapeau-Blondeau, “Suprathreshold stochastic resonance and noise-enhanced Fisher information in arrays of threshold devices,” *Phys. Rev. E*, vol. 68, pp. 1–10, Sep. 2003.
- [12] D. Rousseau, G. V. Anand, and F. Chapeau-Blondeau, “Nonlinear estimation from quantized signals: Quantizer optimization and stochastic resonance,” in *Int. Symp. Physics in Signal and Image Processing*, Grenoble, France, Jan. 2003, pp. 89–92.
- [13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley-Interscience, 2001.
- [14] P. Tichavsky, C. H. Muravchik, and A. Nehorai, “Posterior Cramer–Rao bounds for discrete-time nonlinear filtering,” *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.